

Convex Optimization Approach for Stable Decomposition of Stream of Pulses

Tamir Bendory

Electrical engineering department
Technion - Israel Institute of Technology

Abstract—This paper deals with the problem of estimating the delays and amplitudes of a weighted superposition of pulses, called stream of pulses. This problem is motivated by a variety of applications, such as ultrasound and radar. This paper shows that the recovery error of a tractable convex optimization problem is proportional to the noise level. Additionally, the estimated delays are clustered around the true delays. This holds provided that the pulse meets a few mild localization properties and that a separation condition holds. If the amplitudes are known to be positive, the separation is unnecessary. In this case, the recovery error is proportional to the noise level and depends on the maximal number of delays within a resolution cell.

I. INTRODUCTION

In many engineering and scientific problems, the acquired data is comprised of a weighted super-position of pulses (kernels). Typically, we aim to decompose the stream of pulses into its building blocks, frequently called *atoms*. Ultrasound imaging [27], [28] and Radar [2] function as representative examples. In these applications, a pulse is transmitted and its echoes are reflected from different targets and recorded.

Mathematically, we consider the model

$$y[k] = \sum_{k \in \mathbb{Z}} c_m g_\sigma[k - k_m] + n[k], \quad c_m \in \mathbb{R}, \quad (\text{I.1})$$

where $g_\sigma[k] := g[k/\sigma]$ is a sampled version of the producing kernel $g(t)$ with a sampling spacing of $1/N$, namely $g[k] := g(k/N)$, and $K := \{k_m\}$. We assume that the error term is bounded $\|n\|_1 := \sum_k |n[k]| \leq \delta$, with no additional statistical assumptions. Later on, we will also consider a positive stream of pulses, where the amplitudes are assumed to be positive $c_m > 0$. The acquired data (I.1) can be presented as a spike deconvolution problem, i.e.

$$y[k] = (g_\sigma * x)[k] + n[k],$$

where $'*$ ' denotes a discrete convolution, and

$$x[k] = \sum_{k \in \mathbb{Z}} c_m \delta[k - k_m]. \quad (\text{I.2})$$

The aim of this paper is to suggest a stable approach to decompose the stream of pulses into its atoms by solving a tractable convex optimization program.

A well-known approach to decompose the signal into its atoms is by using parametric methods, such as MUSIC, matrix pencil and ESPRIT [23], [22], [20], [18]. However, these methods tend to be unstable in the presence of noise or model mismatch due to sensitivity of polynomial root

finding. An alternative way is to utilize compressed sensing and sparse representations theorems, relying on the sparsity of the signal (e.g. [14], [17]). However, these fields cannot explain the success of ℓ_1 minimization or greedy algorithms as the dictionaries have high coherence.

Inspired by recent advances in the theory of super-resolution [12], [11], [26], [6], [3], [7], [13], [1], [24], we employed a convex optimization framework based on the existence of interpolating polynomials, frequently called *the dual certificate*. In the next section, we elaborate on the convex optimization framework, and reveal the fundamental conditions, enabling stable decomposition of stream of pulses. Section III presents our main theorems. Particularly, we show that the solution of a convex optimization problem results in a stable and localized decomposition of stream of pulses under a separation condition if the pulse satisfies some mild localization properties. In the non-negative case, i.e. $c_m > 0$, the separation is unnecessary and can be replaced by a weaker condition of Rayleigh regularity. We present all the results and the relevant definitions for a univariate stream of pulses, however we stress that similar results also exist for bivariate stream of pulses. Ultimately, Section IV concludes the work and suggests future extensions.

II. CONVEX OPTIMIZATION APPROACH FOR DECOMPOSITION OF STREAM OF PULSES

In this paper we focus on a convex optimization approach for decomposing a stream of pulses in a noisy environment. We use the Total-Variation (TV) norm as a sparse-promoting regularization. In essence, the TV norm is the generalization of ℓ_1 norm to the real line (for rigorous definition, see for instance [21]). For discrete measures of the form (II.1), we have $\|x\|_{TV} = \sum_m |c_m|$. The framework is based on a duality theorem, frequently called *the dual certificate*, as follows [8]:

Theorem II.1. *Let*

$$x(t) = \sum_m c_m \delta_{t_m}(t), \quad c_m \in \mathbb{R}, \quad T := \{t_m\} \subseteq \mathbb{R}, \quad (\text{II.1})$$

and let $y(t) = \int_{\mathbb{R}} g(t-s) dx(s)$ for a L times differentiable kernel $g(t)$. If for any set $\{v_m\} \in \{-1, 1\}$, there exists a function of the form

$$q(t) = \int_{\mathbb{R}} \sum_{\ell=0}^L g^{(\ell)}(s-t) d\mu_\ell(s), \quad (\text{II.2})$$

for some measures $\{\mu_\ell(t)\}_{\ell=0}^L$, satisfying

$$q(t_m) = v_m, \forall t_m \in T, \\ |q(t)| < 1, \forall t \in \mathbb{R} \setminus T,$$

then x is the unique real Borel measure solving

$$\min_{\tilde{x} \in \mathcal{M}(\mathbb{R})} \|\tilde{x}\|_{TV} \quad \text{subject to} \quad y(t) = \int_{\mathbb{R}} g(t-s) d\tilde{x}(s). \quad (\text{II.3})$$

Proof: Let \hat{x} be a solution of (II.3), and define $\hat{x} = x + h$. The difference measure h can be decomposed relative to $|x|$ as

$$h = h_T + h_{T^c},$$

where h_T is supported in T , and h_{T^c} is supported in T^c (the complementary of T). If $h_T = 0$, then also $h_{T^c} = h = 0$. Otherwise, $\|\hat{x}\|_{TV} > \|x\|_{TV}$ which is a contradiction. If $h_T \neq 0$, we perform a polar decomposition of h_T

$$h_T = |h_T| \text{sgn}(h_T),$$

where $\text{sgn}(h_T)$ is a function on \mathbb{R} with values $\{-1, 1\}$ (see e.g. [21]). By assumption, for any $0 \leq \ell \leq L$

$$\int_{\mathbb{R}} g^{(\ell)}(t-s) d\hat{x}(s) = \int_{\mathbb{R}} g^{(\ell)}(t-s) dx(s),$$

which in turn leads to $\int_{\mathbb{R}} g^{(\ell)}(t-s) dh(s) = 0$. Then, for any q of the form (II.2) we get

$$\begin{aligned} \langle q, h \rangle &= \int_{\mathbb{R}} q(t) dh(t) \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \sum_{\ell=0}^L g^{(\ell)}(s-t) d\mu_\ell(s) \right) dh(t) \\ &= \int_{\mathbb{R}} \sum_{\ell=0}^L d\mu_\ell(s) \underbrace{\int_{\mathbb{R}} g^{(\ell)}(s-t) dh(t)}_0 \\ &= 0. \end{aligned}$$

By assumption, for the choice $v_m = \text{sgn}(h_T(t_m))$, there exists q of the form (II.2), such that

$$q(t_m) = \text{sgn}(h_T(t_m)), \forall t_m \in T, \\ |q(t)| < 1, \forall t \in \mathbb{R} \setminus T.$$

Consequently,

$$0 = \langle q, h \rangle = \langle q, h_T \rangle + \langle q, h_{T^c} \rangle = \|h_T\|_{TV} + \langle q, h_{T^c} \rangle.$$

If $h_{T^c} = 0$, then $\|h_T\|_{TV} = 0$, and $h = 0$. Alternatively, if $h_{T^c} \neq 0$, from the second property of q ,

$$|\langle q, h_{T^c} \rangle| < \|h_{T^c}\|_{TV}.$$

Thus,

$$\|h_{T^c}\|_{TV} > \|h_T\|_{TV}.$$

As a result, using the fact that \hat{x} has minimal TV norm, we get

$$\begin{aligned} \|x\|_{TV} &\geq \|x + h\|_{TV} = \|x + h_T\|_{TV} + \|h_{T^c}\|_{TV} \\ &\geq \|x\|_{TV} - \|h_T\|_{TV} + \|h_{T^c}\|_{TV} > \|x\|_{TV}, \end{aligned}$$

Kernels	Gaussian := $e^{-\frac{t^2}{2}}$	Cauchy := $\frac{1}{1+t^2}$
C_0	1.22	1
C_1	1.59	1
C_2	2.04	2
C_3	2.6	5.22
$g^{(2)}(0)$	-1	-2
empirical ν	1.1	0.45

TABLE I

THE TABLE PRESENTS THE NUMERICAL CONSTANTS OF THE GLOBAL PROPERTY IN DEFINITION III.1 FOR THE GAUSSIAN AND CAUCHY KERNELS. ADDITIONALLY, WE EVALUATED BY NUMERICAL EXPERIMENTS THE MINIMAL EMPIRICAL VALUE OF ν , THE SEPARATION CONSTANT OF DEFINITION III.2 FOR EACH KERNEL.

which is a contradiction. Therefore, $h = 0$, which implies that x is the unique solution of (II.3). ■

In practice, we cannot solve the infinite dimensional convex optimization problem (II.3). Hence, we assume that the signal lies on a grid with spacing of $1/N$ which can be as fine as desired. In this case, the TV minimization (II.3) reduces to standard ℓ_1 minimization that can be solved by many existing solver. The solution of the discrete problem converges to solution on the continuum (in the sense of measures) as the discretization becomes finer [25]. The behaviour of the solution in high SNR regime is analyzed in [15], [16].

III. MAIN RESULTS

As aforementioned, the problem of decomposing a stream of pulses can be reduced to the construction of an interpolating function, comprised of the kernel $g(t)$ and its derivatives. The existence of such function relies on two interrelated pillars. First, the kernel should satisfy some localization properties, as follows:

Definition III.1. A kernel g is admissible if it has the following properties:

- 1) $g \in \mathcal{C}^3(\mathbb{R})$, is real and even.
- 2) Global property: There exist constants $C_\ell > 0, \ell = 0, 1, 2, 3$ such that $|g^{(\ell)}(t)| \leq C_\ell / (1+t^2)$, where $g^{(\ell)}(t)$ denotes the ℓ^{th} derivative of g .
- 3) Local property: There exist constants $\varepsilon, \beta > 0$ such that
 - a) $g(t) > 0$ for all $|t| \leq \varepsilon$ and $g(t) < g(\varepsilon)$ for all $|t| > \varepsilon$,
 - b) $g^{(2)}(t) < -\beta$ for all $|t| \leq \varepsilon$.

Two typical examples for admissible kernels are the Gaussian kernel $g(t) = e^{-\frac{t^2}{2}}$ and the Cauchy kernel $g(t) = \frac{1}{1+t^2}$ as presented in Table I.

The second pillar is a kernel-dependent separation condition, as follows:

Definition III.2. A set of points $K \subset \mathbb{Z}$ is said to satisfy the minimal separation condition for a kernel-dependent $\nu > 0$ and a given $N, \sigma > 0$ if

$$\min_{k_i, k_j \in K, k_i \neq k_j} |k_i - k_j| \geq \nu \sigma N.$$

In [8], [5] we proved that if the kernel g is admissible and the signal's support satisfies the kernel-dependent separation condition, then there exist constants $\{a_m\}$ and $\{b_m\}$ such that a function of the form

$$q(t) = \sum_m a_m g_\sigma(t - t_m) + g_\sigma^{(1)}(t - t_m),$$

satisfies the interpolation requirements of Theorem II.1. Hence, by Theorem II.1 the ℓ_1 minimization (assuming that the signal lies on the grid) recovers x exactly from y . Additionally, the existence of the interpolating function is the key for proving the robustness and localization of the solution in a noisy environment. These results are summarized in the following theorem:

Theorem III.3. *Consider the model (I.1) for an admissible kernel g . Let us denote the solution of*

$$\min_{\tilde{x}} \|\tilde{x}\|_1 \quad \text{subject to} \quad \|y - g_\sigma * \tilde{x}\|_1 \leq \delta,$$

as $\hat{x} = \sum_m \hat{c}_m \delta[k - \hat{k}_m]$ and $\hat{K} := \{\hat{k}_m\}$. If K satisfies the separation condition of Definition III.2 for $N, \sigma > 0$, then (for sufficiently large ν)

$$\|\hat{x} - x\|_1 \leq \frac{16\gamma^2}{\beta} \delta, \quad (\text{III.1})$$

where $\gamma := \max\{N\sigma, \varepsilon^{-1}\}$. Additionally, if $\varepsilon \geq \tilde{\varepsilon} := \sqrt{\frac{g(0)}{C_2 + \beta/4}}$ we have the following localization properties:

- 1) For any $k_m \in K$, if $c_m \geq 2\delta D_2 \left(1 + \max\left\{\frac{1}{D_1 \varepsilon^2}, \frac{4C_2}{(N\sigma)^2 \beta}\right\}\right)$, then there exists $\hat{k}_m \in \hat{K}$ such that

$$\left|k_m - \hat{k}_m\right| \leq N\sigma \sqrt{\frac{2D_2 \delta}{D_1 \left(|c_m| - 2\delta D_2 \left(1 + \max\left\{\frac{1}{D_1 \varepsilon^2}, \frac{4C_2}{(N\sigma)^2 \beta}\right\}\right)\right)}}.$$

2)

$$\sum_{\{\hat{k}_m \in \hat{K} : |\hat{k}_m - k_n| > N\sigma, \forall k_n \in K\}} |\hat{c}_m| \leq \frac{2D_2}{D_1 \varepsilon^2} \delta,$$

where

$$D_1 := \frac{\beta}{4g(0)},$$

$$D_2 := \frac{3\nu^2 \left(3|g^{(2)}(0)|\nu^2 - \pi^2 C_2\right) + \frac{16C_1^2 \gamma^2 \pi^2}{\beta} \left(1 + \frac{\pi^2}{6\nu^2}\right)}{\left(3|g^{(2)}(0)|\nu^2 - \pi^2 C_2\right) \left(3g(0)\nu^2 - 2\pi^2 C_0\right)}.$$

Remark III.4. A tighter estimation of (III.1) can be found in [8].

In many applications, the underlying signal (I.2) is known to be non-negative, i.e. $c_m > 0$. For instance, in single-molecule microscopy we measure the convolution of positive point sources with the microscope's point spread function [19], [9], [10]. It has become evident that in this case the separation is unnecessary and can be replaced by the weaker condition of Rayleigh regularity, defined as follows:

Definition III.5. We say that the set $\mathcal{P} \subset \{k/N\}_{k \in \mathbb{Z}} \subset \mathbb{R}$ is Rayleigh-regular with parameters (d, r) and write $\mathcal{P} \in \mathcal{R}^{idx}(d, r)$ if every interval $(a, b) \subset \mathbb{R}$ of length $\mu(a, b) = d$ contains no more than r elements of \mathcal{P} :

$$|\mathcal{P} \cap (a, b)| \leq r.$$

Equipped with Definition III.5, we state the main theorem for the non-negative case. This result implies that the recovery error is proportional to the noise level δ , and depends exponentially in the signal Rayleigh regularity r .

Theorem III.6. [4] Consider the model (I.1) with $c_m > 0$ for an admissible kernel g satisfying $g(t) \geq 0$. Then, there exists $\nu > 0$ such that if $\text{supp}(x) \in \mathcal{R}^{idx}(\nu\sigma, r)$ and $N\sigma > \left(\frac{1}{2}\right)^{\frac{1}{2r}+1} \sqrt{\frac{\beta}{g(0)}}$, the solution \hat{x} of the convex problem

$$\min_{\tilde{x}} \|\tilde{x}\|_1 \quad \text{subject to} \quad \|y - g_\sigma * \tilde{x}\|_1 \leq \delta, \tilde{x} \geq 0,$$

satisfies (for sufficiently large ν)

$$\|\hat{x} - x\|_1 \leq \frac{2(2^r - 1)}{C_0} \left(\frac{32C_0}{\beta}\right)^r \gamma^{2r} \delta, \quad (\text{III.2})$$

where $\gamma := \max\{N\sigma, \varepsilon^{-1}\}$.

Remark III.7. A tighter estimation of (III.2) can be found in [4].

IV. CONCLUSION

In this work we have shown that a standard convex optimization program can decompose a stream of pulses into its building blocks. In the general case (i.e. $c_m \in \mathbb{R}$), we have shown that the solution is robust in a noisy environment and that its support is clustered around the support of the sought signal. This holds provided that the convolution kernel g is sufficiently localized and a kernel-dependent separation condition holds.

In the non-negative case, we have proven that the separation condition can be replaced by a weaker condition of Rayleigh regularity. The recovery error in this case is proportional to the noise level and depends on the signal's regularity. It is essentially important to derive the localization properties in this case as well.

The model presented in this work suits many practical applications where a signal is observed through a convolution kernel (typically, the point spread function of a sensing device). In previous work [5], we applied our algorithm for estimating the reflectors in *in-vitro* ultrasound experiments. The experiments showed promising results that corroborate our theoretical findings. It is of great interest to examine these theoretical results across more applications, particularly in the field of computational imaging.

ACKNOWLEDGMENT

I would like to express my deep gratitude to Prof. Shai Dekel, Prof. Arie Feuer, Prof. Dan Adam and Avinoam Bar-Zion for the fruitful collaborations.

REFERENCES

- [1] Jean-Marc Azais, Yohann De Castro, and Fabrice Gamboa. Spike detection from inaccurate samplings. *Applied and Computational Harmonic Analysis*, 2014.
- [2] Omer Bar-Ilan and Yonina C. Eldar. Sub-nyquist radar via doppler focusing. *IEEE Transactions on Signal Processing*, 62(7):1796–1811, 2014.
- [3] T. Bendory, S. Dekel, and A. Feuer. Super-resolution on the sphere using convex optimization. *Signal Processing, IEEE Transactions on*, 63(9):2253–2262, May 2015.
- [4] Tamir Bendory. Robust recovery of positive stream of pulses. *submitted*, 2015.
- [5] Tamir Bendory, Avinoam Bar-Zion, Dan Adam, Shai Dekel, and Arie Feuer. Localized recovery of stream of pulses with application to ultrasound imaging. *in final stages of preparation*, 2015.
- [6] Tamir Bendory, Shai Dekel, and Arie Feuer. Exact recovery of dirac ensembles from the projection onto spaces of spherical harmonics. *Constructive Approximation*, pages 1–25, 2013.
- [7] Tamir Bendory, Shai Dekel, and Arie Feuer. Exact recovery of non-uniform splines from the projection onto spaces of algebraic polynomials. *Journal of Approximation Theory*, 182(0):7 – 17, 2014.
- [8] Tamir Bendory, Shai Dekel, and Arie Feuer. Robust recovery of stream of pulses using convex optimization. *submitted*, 2014.
- [9] Eric Betzig, George H Patterson, Rachid Sougrat, O Wolf Lindwasser, Scott Olenych, Juan S Bonifacio, Michael W Davidson, Jennifer Lippincott-Schwartz, and Harald F Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006.
- [10] I Bronstein, Y Israel, E Kepten, S Mai, Yaron Shav-Tal, E Barkai, and Y Garini. Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Physical review letters*, 103(1):018102, 2009.
- [11] Emmanuel J Candès and Carlos Fernandez-Granda. Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254, 2013.
- [12] Emmanuel J Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 2013.
- [13] Yohann De Castro and Fabrice Gamboa. Exact reconstruction using beurling minimal extrapolation. *Journal of Mathematical Analysis and applications*, 395(1):336–354, 2012.
- [14] David L Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.
- [15] Vincent Duval and Gabriel Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, pages 1–41, 2013.
- [16] Vincent Duval and Gabriel Peyré. Sparse spikes deconvolution on thin grids. *arXiv preprint arXiv:1503.08577*, 2015.
- [17] Michael Elad. *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer, 2010.
- [18] Yingbo Hua and Tapan K Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 38(5):814–824, 1990.
- [19] Thomas A Klar, Stefan Jakobs, Marcus Dyba, Alexander Egner, and Stefan W Hell. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proceedings of the National Academy of Sciences*, 97(15):8206–8210, 2000.
- [20] Richard Roy and Thomas Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(7):984–995, 1989.
- [21] Walter Rudin. *Real and complex analysis (3rd)*. New York: McGraw-Hill Inc, 1986.
- [22] Ralph O Schmidt. Multiple emitter location and signal parameter estimation. *Antennas and Propagation, IEEE Transactions on*, 34(3):276–280, 1986.
- [23] Petre Stoica and Randolph L Moses. *Spectral analysis of signals*. Pearson/Prentice Hall Upper Saddle River, NJ, 2005.
- [24] Gongguo Tang. Resolution limits for atomic decompositions via markov-bernstein type inequalities. *preprint*, 2015.
- [25] Gongguo Tang, Badri Narayan Bhaskar, and Benjamin Recht. Sparse recovery over continuous dictionaries-just discretize. In *Signals, Systems and Computers, 2013 Asilomar Conference on*, pages 1043–1047, Nov 2013.
- [26] Gongguo Tang, Badri Narayan Bhaskar, Parikshit Shah, and Benjamin Recht. Compressed sensing off the grid. *Information Theory, IEEE Transactions on*, 59(11):7465–7490, 2013.
- [27] Ronen Tur, Yonina C Eldar, and Zvi Friedman. Innovation rate sampling of pulse streams with application to ultrasound imaging. *Signal Processing, IEEE Transactions on*, 59(4):1827–1842, 2011.
- [28] Noam Wagner, Yonina C Eldar, and Zvi Friedman. Compressed beam-forming in ultrasound imaging. *Signal Processing, IEEE Transactions on*, 60(9):4643–4657, 2012.